DOCUMENT RESUME

ED 105 374 CS 001 680

AUTHOR Layton, James R.
TITLE Diagnostic Reading Evaluations: Which Scores to Report.
PUB DATE Feb 75
NOTE 16p.; Paper presented at the Annual Meeting of the Plains Regional Conference of the International Reading Association (3rd, St. Louis, February 20-22, 1975)

EDRS PRICE MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS Reading; *Reading Ability; *Reading Diagnosis; Reading Improvement; *Reading Research; Reading Skills; *Reading Tests; *Remedial Reading; Teaching Techniques
IDENTIFIERS Slosson Intelligence Test; Spache Diagnostic Reading Scales

ABSTRACT

The purposes of this study were to determine which test instruments should be used to report instructional and independent reading levels to tutors or classroom teachers and what recommendations should be made concerning the materials to be used for teaching the child. The subjects were children from grades two through ten who had been identified as problem readers by their classroom teachers. The students were instructed by college students who were instructed in the administration of the Spache Diagnostic Reading Scales, the Slosson Intelligence Test, and the informal reading inventory. The teachers were also given a knowledge test and a performance test on the administration of the measures. The subjects were instructed for approximately 30 hours during a five-week period. The results led to the following conclusions: the informal reading inventory appears to be the safest and most reasonable device for reporting reading levels to teachers; the Spache Diagnostic Reading Scales did provide for a better analysis of oral reading performance, since it allowed more errors than the informal reading inventory but not enough to produce anxiety; and the Diagnostic Reading Scales measure reading growth accurately if grade equivalents are not considered. (WR)

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

James R. Layton
Springfield, Missouri
February 14, 1975

PERMISSION TO REPRODUCE THIS COPYRIGHTED MATERIAL HAS BEEN GRANTED BY
James R. Layton
TO ERIC AND ORGANIZATIONS OPERATING UNDER AGREEMENTS WITH THE NATIONAL INSTITUTE OF EDUCATION. FURTHER REPRODUCTION OUTSIDE THE ERIC SYSTEM REQUIRES PERMISSION OF THE COPYRIGHT OWNER

ED105374

DIAGNOSTIC READING EVALUATIONS: WHICH SCORES TO REPORT

In 1942, Killgallon devised a set of standards for determining the independent and instructional levels of childrens' reading performance. Betts presented Killgallon's work, the informal reading inventory, as an appropriate evaluation system and it is still utilized today.[1] Although there have been recommended changes to Betts' criteria for the informal reading inventory, few of the scoring alternatives varied significantly from the original.[2] Spache developed a test which drastically lowered both the standards for identifying the two levels of reading performance and the accuracy criteria for both word recognition and comprehension.[3] Durrell's diagnostic test, although requiring a high degree of accuracy, is scored on an additional criteria of rate.[4] This additional element can influence a child's scores significantly. Administering any of the instruments above to the same child, under similar conditions, within a short time range, will result in a wide variety of instructional and independent reading levels. Although the child's reading ability will be the same,

---

[1]Emmett A. Betts, Foundations of American Reading Instruction (New York: American Book Co., 1950) pp. 484-485.

[2]James R. Layton and Bruce Turner, "Some Corrective Procedures for Classroom Reading Problems." Paper read at International Reading Association Convention, New Orleans, May, 1974.

[3]George D. Spache, Diagnostic Reading Scales, Examiner's Manual (Del Monte Park, California: California Test Bureau, 1963) p. 7.

[4]Donald D. Durrell, Durrell Analysis of Reading Difficulty, New York: Harcourt, Brace and Javonovich, 1955).

CS 001 680

the scores will not reflect such. These variations may be difficult to report to a teacher who has the responsibility of teaching the child, although the examiner will understand them quite well. The teacher may experience difficulty also when placing a child in material based on a diagnostician's findings (if any of the foregoing tests are used) due to the variability which is found in published materials.[5] Research indicates that a wide variety of grade level ratings can be found in textbooks and that a large percentage of children are placed in material which is too difficult.[6]

This report presents the conclusions drawn from two major questions which were explored:

1. Which test instruments should be used to report instructional and independent reading levels to the tutor or classroom teacher?
2. What recommendations should be made concerning the materials to be used for teaching the child?

The students in this study were identified as problem readers by classroom teachers and referred to the Summer Reading Program at Southwest Missouri State University. Children from grade two through grade ten were used. Chronological ages ranged from seven to fifteen. Randomization of socio-economic status was established although it was not a part of the study. Only children whose intelligence quotients were above eighty-five, as measured by the Slosson Intelligence Test, were included in the study.

---

[5]James R. Layton, "The Relationship of Science Vocabulary Reading, General Reading, Intelligence, and Social Residence...Students" (unpublished Ed.D. dissertation, University of Southern Mississippi, Mississippi, 1972), pp. 13-14.

[6]Ibid., p. 15-16.

The data were collected during the summers of 1973-74. The second summer was used to confirm the conclusions drawn during the first summer and to extend the nature of the study. Fifty-three public school children were involved the first summer and twenty-seven the second summer. The low number of students the second summer limited the findings. No comparisons between the two summer groups were made. In both studies, the procedures used by the teachers were as identical as conditions would allow. The children received similar amounts of instructional time both summers.

The college students who taught the children had the following characteristics:

1. All held a Bachelor of Science Degree in teaching. The majority of both groups were elementary majors. Eleven percent of thirty-five teachers during the first summer and 7% of twenty-seven teachers the second summer held secondary education degrees.
2. All of the teachers had completed two courses in reading: (1) reading methods and (2) remedial reading techniques.
3. Almost all of the teachers had completed one course in individual testing which required proficiency in the administration of the Peabody Picture Vocabulary Test and the Slosson Intelligence Test.
4. The range of teaching experience was from zero to twenty-four years within the total number of both groups. This variable was not a part of the study.
5. The teachers in both groups spent approximately the same amount of time both in learning and teaching.

To insure that the teachers could conduct the teaching and for control of the teaching to conduct this study, the following procedures were used during both summer sessions:

1. The teachers were instructed in the administration of the Spache Diagnostic Reading Scales and the Slosson Intelligence Test, and the informal reading inventory. Designing learning programs based on the tests' results were learned by all teachers.

2. The teachers had to pass both a knowledge test and a performance test on the administration of the tests to be used and procedures to be followed.

Fifteen hours of directed activity was scheduled for the teachers the first week for learning to administer the tests and for learning the procedures to be followed in teaching. The children arrived the second week. Other procedures were studied during the first week but they are not specifically related to this study.

Identical teaching schedules were used during both summers. Logically, flexibility in procedures were allowed to care for individual differences. The children in both summer sessions received approximately thirty hours of instruction in a five week period (this excludes pre-, and post-testing). The one-hour daily period was divided as follows:

Phase I - Directed teaching of appropriate skills.

Thirty Minutes
Student Centered

A. Word identification and recognition.

B. Vocabulary (meaning).

C. Comprehension (sentence and paragraph).

D. Mechanics.

Fifteen Minutes
Student Centered

A. Additional testing and evaluation.

B. Recreational reading.

C. Worthwhile activities.

Phase III - Culminating activities.

Fifteen Minutes
Teacher Centered

A. Putting away materials.

B. Getting ready for the next child.

C. Sketching the next lesson.

After the children's appropriate instructional and independent reading levels were established and skills needs identified, the teachers followed the time sequence above. A variety of materials were provided. In addition to published materials, the teachers were instructed in language experience techniques for teaching and in developing individual, creative progress charts for motivational and learning purposes. Both language experience activities and progress charts were designed to fit the child's major interests. The progress charts were used to exhibit gain in any specific area the child wished to measure. The progress charts were especially successful.

The teachers were required to spend two individualized, instructional hours daily in class without the children. Instructions in various aspects of remedial reading were given. Procedures and techniques for determining readability levels and writing diagnostic reports were presented.

Problems arose when children could not read materials which had publishers' designations the same as the independent and instructional levels ascertained through the Diagnostic Reading Scales. Variations existed when the childrens' levels on the Diagnostic Reading Scales were compared to the informal reading inventory's levels. Additionally, a wide

variation existed within the materials to be used by the children. All of these discoveries were made by the practicum teachers during the first summer. Their questions, which appear below, were responsible for this study.

1. Is there a difference in the readability levels of the reading selections contained in the Diagnostic Reading Scales and a published informal reading inventory?
2. Are the independent and instructional levels of the Diagnostic Reading Scales or the informal reading inventory more useful to a classroom teacher who will receive the final, written evaluation?
3. After a choice is made as to the test results to report how can a teacher be assured that the results reported will correlate with the materials used by the teacher?

This study was done at the end of both summers. The collection and analysis of the data led to the following conclusions:

1. The readability levels of the selections within both the Diagnostic Reading Scales and the informal reading inventory as judged through application of the Fry Readability Formula[7] are very similar at each level. There was an indication that a few of the upper level selections of the Diagnostic Reading Scales were out of progression.
2. The claim by Spache that children comprehend at a higher level in silent reading than in oral reading, using the directions of the Diagnostic Reading Scales was supported. The children's mean

---

[7]Edward Fry, "A Readability Formula That Saves Time," Journal of Reading, XI (April, 1968), 513-516.

grade equivalent during the first summer on word recognition was lower than oral reading. Oral reading mean scores were lower than silent reading means. Listening level means were the highest. Correlations between the subtests were computed.

PEARSON PRODUCT MOMENT
CORRELATIONS BETWEEN SUBTESTS ON THE
SPACHE DIAGNOSTIC READING SCALES

N=53 First Summer

	Word Recognition	Oral Reading	Silent Reading	Listening
Word Recognition	-	.869	.877	.643
Oral Reading	-	-	.909	.674
Silent Reading	-	-	-	.817
Listening	-	-	-	-

The word recognition and selection reading correlations are extremely strong. The listening level does not strongly correlate with word recognition or oral reading but is highly correlated with silent reading. Computation of expectancy levels for the fifty-three subjects using the Bond-Clymer formula[8] and subsequent correlation with listening level as determined by the Diagnostic Reading Scales yielded an $r = .791$. No attempt was made to see which was the more accurate scale. It was concluded that the children's scores on listening were correlated with intelligence scores. The data presented above was repeated the second summer and found similar.

[8]Guy L. Bond and Miles A. Tinker, Reading Difficulties: Their Diagnosis and Correction (New York: Appleton-Century-Crofts, 1967), p. 92.

A serious problem arose when attempts were made to place children in publishers' materials based on the scores of the Diagnostic Reading Scales. Assuming the results of tests to be accurate, readability levels for published materials were determined. Fry's formula was utilized.[9] This procedure to place children in material at the levels identified did not result in success. Even though variability was found in the materials, when the levels of the Diagnostic Reading Scales were used, the children exhibited extreme difficulty in mechanics and word recognition during reading and in specific comprehension skills in silent reading. The consensus of the teachers was that the children were overplaced. During the first summer the situation was rectified by placing the students in materials using the informal reading inventory criteria and graded texts according to Betts' suggestions.

Although the Diagnostic Reading Scales did not provide accurate levels for placement, it was used as a post-test measure. Using the scales' criteria, the students in the summer 1973 program, showed significant progress in oral (eight months gain, $t = 2.3569 > .05$) and silent (eight months gain, $t = 2.2812 > .05$) reading. Gains of five months on word recognition was not significant. The gains were measured after thirty hours of instruction.

During the second summer, the same procedures were followed with similar results. Additional data were collected. The Diagnostic Reading Scales were administered in a manner that would allow for scoring by both

---

[9]Fry, "A Readability Formula That Saves Time," p. 92.

Spache's criteria and by Betts' criteria for the informal reading inventory. Turner's informal reading inventory[10] and the Diagnostic Reading Scales were regraded by Fry's readability formula and determined similar enough to subject scores made on them to comparison. Due to the opposing views of Betts and Spache toward instructional and independent reading levels and the yield of only a comprehension score in silent reading, only oral reading (word recognition, mechanics, and comprehension) performance scores were used in the study. The results of the comparisons and correlations were limited by the small number used. The comparisons yielded similar standard deviations, but very dissimilar means and ranges. Consistently, on group and individual comparisons, the Diagnostic Reading Scales (scored according to Spache's criteria) yielded higher scores than other measures. Product-moment correlations among the variables yielded coefficients which were significant at the .001 level.

MEANS, RANGES, AND STANDARD DEVIATIONS FOR THE INFORMAL READING INVENTORY, AND DIAGNOSTIC READING SCALES

Test	N	Grade Equivalent Range	Mean	S.D.
I.R.I.	27	.5-6.0	2.25	1.56
D.R.S. (Spache)	27	1.6-7.5	4.20	1.67
D.R.S. (Betts)	27	.8-7.5	3.57	1.80

[10]Bruce Turner, Informal Reading Inventory (Mimeographed test; Springfield, Missouri: Center for Reading, Southwest Missouri State University, 1972).

PRODUCT-MOMENT
CORRELATIONS OF SCORES OBTAINED ON
THE DIAGNOSTIC READING SCALES AND
THE INFORMAL READING INVENTORY

N	Variables	Diagnostic Reading Scales (Spache)	Diagnostic Reading Scales (Betts)
27	Informal Reading Inventory	.749	.746

A t test comparison was made of the scores on the three variables. When the students' informal reading inventory scores were compared to their Diagnostic Reading Scales scores first by Betts' criteria; then Spache's criteria, significant differences were found at the .001 level. Although the three variables showed strong correlations, the Diagnostic Reading Scales, even with altered scoring standards, continued to score significantly higher than the informal reading inventory.[11]

t TEST BETWEEN THE INFORMAL
READING INVENTORY AND DIAGNOSTIC READING SCALES

N	Variables	D.R.S. (Spache)	D.R.S. (Betts)
27	I.R.I.	t=8.12	t=5.50

The consensus of the teachers who administered the test was that the procedures recommended by Betts were more appropriate for placing children in materials. This decision was based on the responses of the students to the easier levels as identified by the informal reading inventory and the gains made by the students.

[11] Denise N. Huddleston. "A Comparison of Two Standards Applied to the Instructional Reading Level. . .As Determined by Classroom Teachers." (unpublished seminar paper, Southwest Missouri State University, 1974), p. 26.

The data presented were analyzed and the following conclusions drawn:

1. Test instruments used to report independent and instructional reading levels to tutors or classroom teachers should be developed to reflect reading levels commensurate with the materials being used in the classroom. Reading levels should be reported in such a manner as to insure that an unknowing teacher will not overplace a child, continue his frustration, and/or thwart his progress.

2. The informal inventory based on the classroom materials and administered according to Betts' criteria appears to be the safest and most reasonable device for reporting reading levels to teachers. The incident of the independent level being higher than the instructional level was infrequent in classroom teacher judgment and tests except the Diagnostic Reading Scales.

3. There are children who read poorly orally due to anxiety and silent reading comprehension becomes superior to oral as a child progresses through the grades. Nevertheless, for children in this study identified as problem readers, strict application of Spache's standards when reporting test results outside a clinic or special reading class may have been harmful.

4. The Diagnostic Reading Scales did provide for a better analysis of oral reading performance since it allowed more errors than the informal reading inventory but not enough to produce anxiety. The Diagnostic Reading Scales is a superior instrument to use as a guide to observe, collect, analyze, and report weaknesses related to the act of reading. It is suggested that it be used in teacher training institutions and inservice training.

5. The Diagnostic Reading Scales will measure accurately the growth that children make in all phases of reading if grade equivalents are not considered.

6. School systems, clinicians, and libraries in a given area should establish standards for evaluating the readability levels of books and materials. One standard should be used. As we establish standards for student evaluation and text evaluation, it should be easier to match a child with a book.

7. Efforts should be made in teacher preparation institutions and in inservice training to develop the teachers' skill in diagnosis. Clinics cannot properly do this for all the children who need it.

8. Studies should also be conducted to determine the relationship between various readability measures. This would serve to eliminate determining the readability for all incoming materials. A knowledge of readability measurement techniques will also aid the teacher in writing or developing materials and/or measuring the written compositions of the students.

In summary, this study was not intended to make recommendations concerning all the aspects of reporting the results of reading tests. Many data devices must be used. Farr wrote:

> Estimates of student performance based on classroom materials is probably of greatest value to the classroom teacher and the IRI can provide this information. . .but its value decreases in the upper grades.[12]

---

[12] Roger Farr, Reading: What Can Be Measured. (Newark: International Reading Association, 1969).

Fry supports the use of providing comfortable materials for children in school;[13] the informal reading inventory can aid the teachers' task in providing this type of school atmosphere. Harris asks, "How difficult a book can this child read?"[14]

This paper intended to offer evidence that there is a way to find the appropriate level of reading to insure a child's proper placement in reading materials. Although we sometimes get low scores on the I.R.I., moving up is hardly a problem. It is a problem when you begin too high and further frustrate a frustrated child; then begin to move backward. Why not start right the first time?

---

[13]Edward Fry, Reading Instruction for Classroom and Clinic (New York: McGraw-Hill Book Company, 1972), p. 14.

[14]Albert J. Harris, How to Increase Reading Ability (New York: David McKay, 1970), p. 139.

James R. Layton
Springfield, Missouri
February 14, 1975

BIBLIOGRAPHY

Betts, Emmett A. Foundations of American Reading Instruction. American Book Company, New York, 1950.

Bond, Guy L. and Miles A. Tinker. Reading Difficulties: Their Diagnosis and Correction. Appleton-Century-Crofts, New York, 1967.

Durrell, Donald D. Durrell Analysis of Reading Difficulty. Harcourt, Brace and Javonovich, New York, 1955.

Farr, Roger. Reading: What Can Be Measured. International Reading Association, 1969.

Fry, Edward. "A Readability Formula That Saves Time," Journal of Reading, XI, April, 1968, pp. 513-516.

_____. Reading Instruction for Classroom and Clinic. McGraw-Hill Book Company, New York, 1972.

Harris, Albert J. How to Increase Reading Ability. David McKay, New York, 1970.

Huddleston, Denise N. "A Comparison of Two Standards Applied to the Instructional Reading Level. . .As Determined by Classroom Teachers." (unpublished seminar paper, Southwest Missouri State University), 1974.

Layton, James R. "The Relationship of Science Vocabulary Reading, General Reading, Intelligence, and Social Residence. . .Students." (unpublished Ed.D. dissertation, University of Southern Mississippi, Mississippi), 1972.

Layton, James R. and Bruce Turner. "Some Corrective Procedures for Classroom Reading Problems." Paper read at International Reading Association Convention, New Orleans, May, 1974.

Spache, George D. Diagnostic Reading Scales. Examiner's Manual. Del Monte Park, California, 1963.

Turner, Bruce. Informal Reading Inventory. (mimeographed test) Center for Reading, Southwest Missouri State University, Springfield, Missouri, 1972.

James R. Layton
Springfield, Missouri
February 14, 1975

BIBLIOGRAPHY

Betts, Emmett A. Foundations of American Reading Instruction. American Book Company, New York, 1950.

Bond, Guy L. and Miles A. Tinker. Reading Difficulties: Their Diagnosis and Correction. Appleton-Century-Crofts, New York, 1967.

Durrell, Donald D. Durrell Analysis of Reading Difficulty. Harcourt, Brace and Jovonovich, New York, 1955.

Farr, Roger. Reading: What Can Be Measured. International Reading Association, 1969.

Fry, Edward. "A Readability Formula That Saves Time." Journal of Reading, XI, April, 1968, pp. 513-516.

_____. Reading Instruction for Classroom and Clinic. McGraw-Hill Book Company, New York, 1972.

Harris, Albert J. How to Increase Reading Ability. David McKay, New York, 1970.

Huddleston, Denise N. "A Comparison of Two Standards Applied to the Instructional Reading Level. . .As Determined by Classroom Teachers." (unpublished seminar paper, Southwest Missouri State University), 1974.

Layton, James R. "The Relationship of Science Vocabulary Reading, General Reading, Intelligence, and Social Residence. . .Students." (unpublished Ed.D. dissertation, University of Southern Mississippi, Mississippi), 1972.

Layton, James R. and Bruce Turner. "Some Corrective Procedures for Classroom Reading Problems." Paper read at International Reading Association Convention, New Orleans, May, 1974.

Spache, George D. Diagnostic Reading Scales. Examiner's Manual. Del Monte Park, California, 1963.

Turner, Bruce. Informal Reading Inventory. (mimeographed test) Center for Reading, Southwest Missouri State University, Springfield, Missouri, 1972.